

XXXI

jornadas de Automática

Jaén 8 al 10 de septiembre de 2010



UNIVERSIDAD DE JAÉN

ISBN: 978-84-693-0715-1



CEA

[Comunicaciones](#)

[Comités](#)

[Programa](#)

[Patrocinadores](#)



CONSTRUCCIÓN DE MAPAS VISUALES Y LOCALIZACIÓN MEDIANTE MÉTODOS BASADOS EN APARIENCIA GLOBAL

Francisco Amorós, Óscar Reinoso, Luis Payá, Arturo Gil, Jose M^a Marín
Departamento de Ingeniería de Sistemas Industriales,
Universidad Miguel Hernández. Avda. de la Universidad s/n, 03202 Elche (Alicante)
o.reinoso@umh.es, lpaya@umh.es

Resumen

Las técnicas basadas en apariencia global han demostrado ser métodos robustos a la hora de construir mapas topológicos de un entorno usando únicamente información visual. En este documento se propone un mecanismo para construir mapas basados en apariencia global a partir de un conjunto de imágenes omnidireccionales tomadas por un robot a lo largo de un entorno. Para extraer la información más relevante de esas imágenes, se usan distintos métodos de compresión, cuyo comportamiento es analizado y comparado en este trabajo. En este análisis se incluye la invariancia ante rotaciones del robot en el plano del suelo y pequeños cambios en el entorno. El objetivo principal consiste en construir un mapa que el robot pueda usar en cualquier aplicación donde se necesite conocer su posición y orientación dentro de dicho entorno, con el mínimo requerimiento de memoria y coste computacional pero sin renunciar a una precisión razonable. De esta manera, se presentan tanto métodos para construcción de los mapas como los resultados de su funcionamiento para futuras aplicaciones.

Palabras clave: Creación de mapas, métodos basados en apariencia, visión omnidireccional, localización espacial.

1. INTRODUCCIÓN

Las aplicaciones que requieren la navegación de un robot o un equipo de robots a través de un entorno necesita el uso de una representación interna del mismo. Gracias a ésta, el robot puede estimar su posición y orientación dentro de dicho mapa con la información capturada por los diferentes sensores con los que está equipado. De entre los diversos tipos de sensores que existen para llevar a cabo esta tarea, destacan los sistemas de visión omnidireccional debido a la riqueza de la información que proporcionan y su relativo bajo coste. Los trabajos clásicos en robótica móvil basados en visión por computador se han centrado en la extracción de marcas o puntos característicos de la imagen para construir el mapa y llevar a ca-

bo la localización del robot [1], [10], [11], [14], [15]. Sin embargo, no es necesario extraer ese tipo de marcas visuales para reconocer dónde se encuentra el robot. En lugar de eso, se puede procesar la imagen en su conjunto. Estos métodos basados en apariencia global suponen una opción interesante cuando se trabaja con espacios desestructurados donde puede ser difícil encontrar patrones o puntos característicos para reconocer la escena, tal y como se puede ver en [6].

Con estos métodos basados en la apariencia global, las imágenes son almacenadas sin extraer ningún punto, y la comparación se realiza directamente, trabajando con el conjunto total de la información de las escenas. Por tanto, el problema de encontrar la posición del robot en el entorno se reduce al problema de obtener la mejor correspondencia entre la imagen actual y las del mapa de referencia. Como desventaja, se trabaja con una gran cantidad de información, teniendo un alto coste computacional. Ésta es la razón por la que se necesita estudiar técnicas de compresión. Existen numerosos investigadores que muestran técnicas de compresión que pueden ser usadas. Por ejemplo, PCA (Análisis de Componentes Principales) es un método muy extendido que ha demostrado ser robusto en su aplicación al procesamiento de imágenes [6], [13]. Debido al hecho de que PCA convencional no es un método invariante a rotación, otros autores han introducido técnicas que, a pesar de ser más complejas y computacionalmente más pesadas, permiten tener en cuenta imágenes con otras orientaciones además de las almacenadas [4]. También hay autores que utilizan la Transformada de Fourier como un método genérico para extraer la información más relevante de la imagen. En este campo, [7] extiende el concepto de Transformada de Fourier, definiendo la Firma de Fourier, que está basada en la Transformada de Fourier Discreta unidimensional de las filas de la imagen, obteniendo mayor robustez en cuanto a imágenes con distintas rotaciones. Por otro lado, [2] usa un método basado en el Histograma de Gradientes Orientados (HOG) aplicado a la detección de peatones, probando que puede ser un descriptor útil para la visión por computados y el procesamiento de imágenes usando la aparien-

cia de los objetos. En [12] se presenta un método para construir un descriptor de la escena en base al análisis multiescala de diversas características visuales. Por su parte, [9] presenta un estudio comparativo sobre diversos métodos basados en apariencia.

Con el presente artículo se pretende extender este estudio comparativo, teniendo en cuenta tres métodos distintos: Firma de Fourier, PCA aplicado a la Firma de Fourier, y HOG. El resto del artículo se organiza de la siguiente manera: en la sección 2, se presentan las técnicas comentadas. En el siguiente punto se describe la metodología seguida para la construcción del mapa, y se presentan resultados previos en cuanto al coste computacional de dichos procesos. La sección 4 trata sobre la forma en la que se estima la pose del robot en el seno del mapa ya construido, y el tiempo y coste computacional de cada algoritmo. Para finalizar, se realiza un resumen del trabajo con las principales conclusiones.

2. REVISIÓN DE TÉCNICAS DE COMPRESIÓN

En esta sección se resumen algunas técnicas para extraer la información más relevante de una base de datos formada por imágenes panorámicas, tratando de mantener al mínimo la cantidad de información.

2.1. FIRMA DE FOURIER

Tal y como se muestra en [9], es posible representar una imagen usando la Transformada Discreta de Fourier de cada línea. Por ello, se puede expandir cada fila de la imagen $\{a_n\} = \{a_0, a_1, \dots, a_{N-1}\}$ en la secuencia de números complejos $\{A_n\} = \{A_0, A_1, \dots, A_{N-1}\}$:

$$\{A_n\} = \mathcal{F}[\{a_n\}] = \sum_{k=0}^{N-1} a_n e^{-j \frac{2\pi}{N} kn}; k = 0, \dots, N-1 \quad (1)$$

Aprovechando las propiedades de la Transformada de Fourier, tomamos únicamente los primeros coeficientes de la serie para representar cada fila, ya que la información más relevante se concentra en las componentes de baja frecuencia. Además, como estamos trabajando con imágenes omnidireccionales, cuando calculamos la transformada de cada fila tenemos otra propiedad muy interesante: invariancia rotacional. Debido a que la rotación de una imagen panorámica se representa como un desplazamiento de sus columnas, es decir, un desplazamiento a lo largo del eje horizontal de cada fila, el módulo de los coeficientes de las se-

ries de Fourier será el mismo. Lo único que cambiará será la fase de esos coeficientes. Representando cada fila de la imagen original como $\mathcal{F}[\{a_n\}]$ y $\mathcal{F}[\{a_{n-q}\}]$ la misma fila desplazada q píxeles, el teorema del desplazamiento establece que:

$$\mathcal{F}[\{a_{n-q}\}] = A_k e^{-j \frac{2\pi q k}{N}}; \quad k = 0, \dots, N-1 \quad (2)$$

siendo q el desplazamiento proporcional al giro entre imágenes, $\mathcal{F}[\{a_{n-q}\}]$ la Transformada de la secuencia desplazada, y A_k los componentes de la Transformada de la secuencia original. Por lo tanto, podemos hallar la rotación relativa de dos imágenes a través de la comparación de las fases de sus coeficientes de Fourier.

2.2. PCA SOBRE LA FIRMA DE FOURIER

Las técnicas basadas en PCA han demostrado ser un método de compresión muy útil. Hace posible que, teniendo un conjunto de N imágenes con M píxeles cada una, $\vec{x}^j \in \mathbb{R}^{M \times 1}, j = 1 \dots N$, se pueda transformar cada imagen en un vector de características (también llamado proyección de la imagen) $\vec{p}^j \in \mathbb{R}^{k \times 1}, j = 1 \dots N$, siendo $k \leq N$ las características PCA que contienen la información más relevante de la imagen [5]. Sin embargo, si aplicamos PCA directamente sobre las imágenes originales, se obtiene un conjunto de datos con información relativa únicamente a la orientación del robot cuando adquirió las imágenes, pero no para otras posibles orientaciones. Lo que se propone en este punto es transformar los módulos de la Firma de Fourier en lugar de la imagen, obteniendo la compresión de información invariante ante rotación. Con ello se unen las ventajas de las técnicas de PCA y de Fourier.

2.3. HISTOGRAMA DE ORIENTACIÓN DEL GRADIENTE

Los descriptores basados en el Histograma de Orientación del Gradiente (HOG) [2] se basan en la orientación del gradiente en áreas locales de una imagen. El primer paso para aplicar HOG a la imagen es calcular las derivadas espaciales de la imagen a lo largo del eje x e y (I_x y I_y). En este caso, dichas derivadas se han obtenido calculando la convolución de las imágenes con filtros Gausianos de diferente varianza. Una vez que se han calculado las convoluciones, se hallan la magnitud y dirección de los valores del gradiente en cada pixel:

$$|G| = \sqrt{I_x^2 + I_y^2}; \quad \theta = \arctan\left(\frac{I_x}{I_y}\right) \quad (3)$$

Posteriormente, la imagen se divide en celdas. Después se calculan los histogramas de dichas celdas.

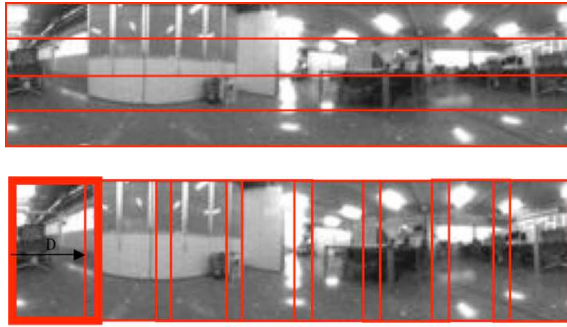


Figura 1: Celdas usadas para calcular la posición actual y la orientación del robot.

Cada píxel de la celda tiene un cierto peso en el histograma de orientación basado en el valor calculado de la magnitud de su gradiente. El número de divisiones del histograma en los experimentos es 8, y el ángulo varía entre -90° y 90° . Cada imagen se representa a través del histograma de cada celda ordenados en un vector. Hay que tener en cuenta que una imagen omnidireccional contiene los mismos píxeles en una fila aunque la imagen esté rotada. Por ello, para llevar a cabo la localización del robot se han tomado las celdas mostradas en la fig. 1(a). De este modo, se obtiene un array de características invariantes ante rotación.

Sin embargo, para conocer la orientación relativa entre dos imágenes rotadas, se usan ventanas verticales con la misma altura que la imagen, pudiendo variar su anchura y la distancia a la que se aplican (fig. 1(b)). Rotando el orden de los histogramas de esas ventanas, se obtiene el mismo resultado que el de calcular el histograma de una imagen rotada un ángulo proporcional a la distancia D . La precisión angular del algoritmo depende de dicha distancia:

$$\text{Angulo}(\circ) = \frac{D * 360}{\text{Anchura de la Imagen}} \quad (4)$$

3. CONSTRUCCIÓN DEL MAPA

En esta sección, se va a evaluar los diferentes métodos con una base de datos formada por imágenes que fueron tomadas en varios espacios bajo condiciones de iluminación reales. Pertenecen a la Facultad Técnica de la Universidad de Bielefeld [8]. Han sido adquiridas con una cámara omnidireccional, y posteriormente convertidas a panorámicas con tamaño 41×256 píxeles. Todas las imágenes pertenecen a espacios interiores. Específicamente, hay ejemplos de una sala de estar (S.E.), una cocina (C.) y un espacio combinado con una cocina y sala de estar (A.C.), todas ellas estructuradas en una cuadrícula o rejilla rectangular de tamaño 10×10 cm. La fig.2 muestra un ejemplo de

Tabla 1: Tamaño del mapa y número de imágenes seleccionadas dependiendo de la cuadrícula.

	Size	10x10	20x20	30x30	40x40
S.E.	22x11	242	66	32	18
C.	12x9	108	30	12	9
A.C.	36x11	396	118	48	27
TOT		746	204	92	54

imágenes correspondientes a cada área. El número de imágenes que compone el mapa varía dependiendo del experimento, ya que, para evaluar la robustez de los algoritmos, se modifica el tamaño de la cuadrícula con la que se seleccionan las imágenes. Así pues, tomaremos cuadrículas de 10×10 cm, 20×20 cm, 30×30 cm y 40×40 cm. En la tabla 1 se muestra el número de imágenes que forman el mapa del robot con respecto la cuadrícula que elegimos.



Figura 2: Ejemplo de imagen de cada zona.

Una vez que se han seleccionado las imágenes de la base de datos, se aplica la técnica de compresión para obtener el conjunto de datos que representa el mapa del robot. Los programas y las simulaciones han sido realizadas usando Matlab R2008b bajo Mac OS X.

La fig.3 muestra la cantidad de memoria y el tiempo necesario para construir el mapa, dependiendo del número de imágenes y de los diferentes parámetros que afectan a cada algoritmo. Estos parámetros son, en el caso de la Firma de Fourier, el número de coeficientes que se toman de cada fila (correspondientes con las frecuencias más bajas). En PCA sobre la Firma de Fourier, se puede variar tanto el número de vectores propios seleccionados de la descomposición PCA como el número de coeficientes de Fourier sobre los que se aplica PCA. Para no tener un exceso de información, se ha fijado el tamaño de la cuadrícula de imágenes que forman el mapa en 20×20 cm. En cuanto a HOG, se puede decidir tanto el tamaño de las celdas como la distancia de aplicación en el caso de las verticales. Como el cálculo de posición y orientación

utilizan celdas totalmente distintas, se separa su estudio de tiempo y memoria.

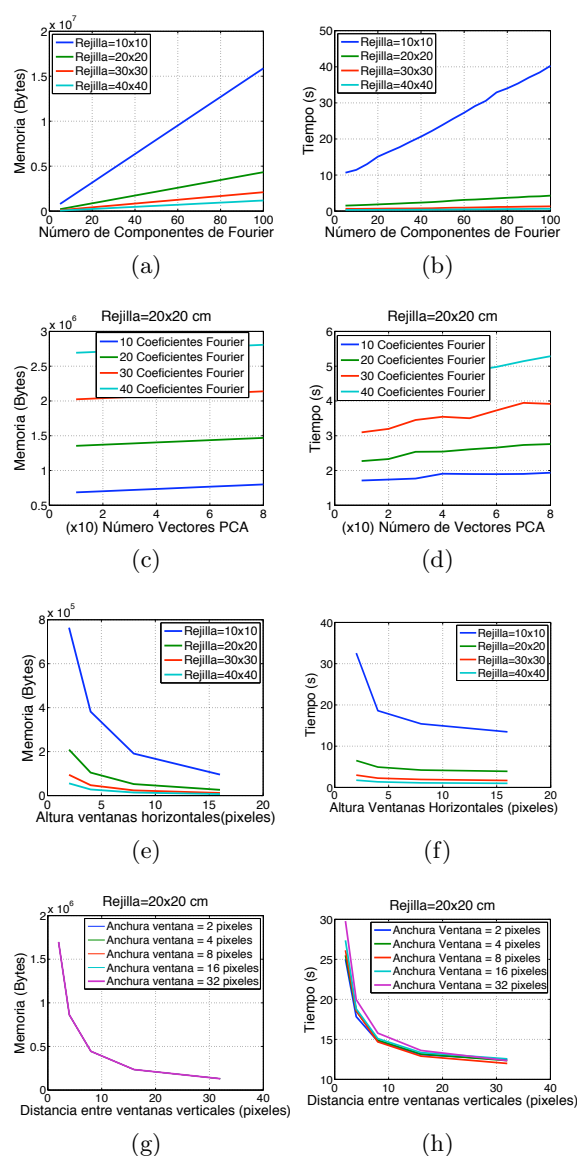


Figura 3: (a) Tiempo empleado y (b) cantidad de memoria requerida para construir el mapa con el algoritmo basado en la Firma de Fourier. (c) Tiempo empleado y (d) cantidad de memoria empleada para construir el mapa usando el algoritmo de PCA sobre la Firma de Fourier. (e) Tiempo empleado y (f) cantidad de memoria para construir el mapa con HOG variando los parámetros de localización. (g) Tiempo empleado y (h) cantidad de memoria para construir el mapa con HOG variando los parámetros de orientación.

En la figura 3(a) y 3(b) se puede ver el tiempo y la memoria necesarias para comprimir y almacenar el mapa con la Firma de Fourier. Naturalmente, existe un incremento proporcional de tiempo y memoria requerida cuando tomamos más coeficientes de cada fila ya que se procesa y almacena más información. Puede ser destacado también los cambios cuando se usan diferentes cuadrículas,

pues el número de imágenes tiene un notable descenso cuando tomamos aquellas más grandes. Las siguientes dos gráficas, las fig. 3(c) y 3(d) muestran los resultados de aplicar PCA sobre la Firma de Fourier. Se puede resaltar la mayor dependencia de los resultados con respecto al número de coeficientes de Fourier que tomamos que con respecto los vectores PCA que se utilizan. En todos los experimentos, las necesidades de tiempo son mayores que en el primer método, pues además de calcular la Firma de Fourier, se debe calcular la descomposición PCA. Las siguientes cuatro figuras estudian la compresión del mapa con HOG. Las fig. 3(e) y 3(f) se obtienen fijando los parámetros de orientación, siendo la anchura de la celda y la distancia de aplicación de las ventanas verticales igual a 8 píxeles (lo que significa un paso en grados de $11,25^\circ$ sin superposición). Estas gráficas muestran una fuerte relación entre el tamaño de la ventana horizontal y la memoria y tiempo requeridos ya que cuanto mayor es la celda, menos se necesitan, pues no se van a superponer. Sin embargo, cuando se estudia el caso de los parámetros de orientación, puede aparecer superposición de las celdas verticales. Por ello, se varía tanto la anchura como la distancia de aplicación de las mismas (fig. 3(g) y 3(h)). En esas gráficas se puede apreciar que el número de ventanas que se utilizan es mucho más influyente en los resultados que su tamaño. La razón es que calcular el histograma de una celda nueva tiene un coste computacional mucho mayor que calcular el histograma de una celda de mayor tamaño. En la fig. 3(g), las líneas de los diferentes casos del gráfico están superpuestas porque la información almacenada, el histograma de cada celda, tiene el mismo tamaño independientemente del tamaño de la celda. Comparando los diferentes métodos, la Firma de Fourier presenta una gran sensibilidad con respecto a la modificación del número de componentes que se toman, pero si se tiene en cuenta que únicamente se almacenan los primeros componentes, los requerimientos de memoria no son muy diferentes al caso en que se aplica la descomposición de los componentes principales a la Firma de Fourier. En cuanto a HOG, la precisión en el cálculo de la orientación (que determina el paso entre ventanas verticales) hará que se tengan mayores o menores requerimientos de tiempo y memoria que en los métodos anteriores.

4. CÁLCULO DE LA POSICIÓN Y ORIENTACIÓN

En esta sección se mide la bondad de cada método a través de la evaluación de los resultados del cálculo de la pose del robot comparando una imagen de entrada nueva con el mapa creado pre-

viamente. Las imágenes de test usadas son todas las que se encuentran en la base de datos, independientemente de la cuadrícula seleccionada para formar el mapa, y 15 rotaciones artificiales de cada una (cada 22.5°). Luego el total de imágenes de test es 11.936. Como la pose incluye la posición y orientación del robot, se lleva a cabo el estudio de ambos, pero por separado. La posición se evalúa como un resultado binario, considerando si se obtiene la mejor correspondencia posible o no, y la información se muestra con medidas Recall-Precision [3]. Cada gráfica tiene información sobre si la localización correcta está en el vecino más cercano (N.N.), es decir, el primer resultado seleccionado, o entre el segundo o tercer vecino más cercano (S.N.N o T.N.N.). El tiempo empleado para calcular la pose del robot se muestra en tablas, medido en segundos. Con respecto a la rotación, se muestra resultados de su precisión en gráfica de barras, dependiendo de cuánto difiere el cálculo del desfase correcto entre la imagen del mapa y la actual. Si el error del experimento es mayor de ± 10 grados, se considera como fallo y no se toma en cuenta.

4.1. FIRMA DE FOURIER

El mapa obtenido con la Firma de Fourier se representa con dos matrices: la de módulos y la de fases de los coeficientes de Fourier. Cuando llega una nueva imagen, el primer paso es calcular su Firma de Fourier. Luego, la localización se estima calculando la distancia Euclídea del espectro de potencia de esa imagen con los espectros del mapa almacenado. La mejor correspondencia es aquella que presenta menor distancia. Una vez se conoce la posición del robot, se utiliza el vector de fases asociado a la imagen más similar para calcular la orientación del robot con respecto al mapa. La tabla 2 muestra el tiempo empleado dependiendo de la cuadrícula de imágenes elegida y el número de coeficientes escogidos de cada fila. Se puede apreciar una mayor dependencia con respecto al número de componentes que con respecto al número de imágenes seleccionadas del mapa para formar la base. Esto es debido a que la estimación de la orientación es la parte que mayor peso tiene en el algoritmo, y depende únicamente del número de componentes de Fourier que escogemos.

La fig. 4 muestra las mediciones de Recall y Precision. Se comprueba que cuanto más coeficientes tomamos, aparece una mejor localización, pero existe un límite a partir del cual no es interesante incrementar el número de elementos que tomamos, pues los resultados no mejoran. En este caso, con 10 componentes se tiene una localización satisfactoria. La precisión en la fase también mejora cuando se toman más coeficientes, aunque se man-

Tabla 2: Tiempo empleado en la estimación de la pose empleando la Firma de Fourier variando el número de coeficientes, en segundos.

GRID	2Com	5Com	10Com	20Com
10x10	0.0099	0.0130	0.0219	0.0487
20x20	0.0072	0.0107	0.0192	0.0446
30x30	0.0070	0.0104	0.0187	0.0440
40x40	0.0069	0.0103	0.0185	0.0438

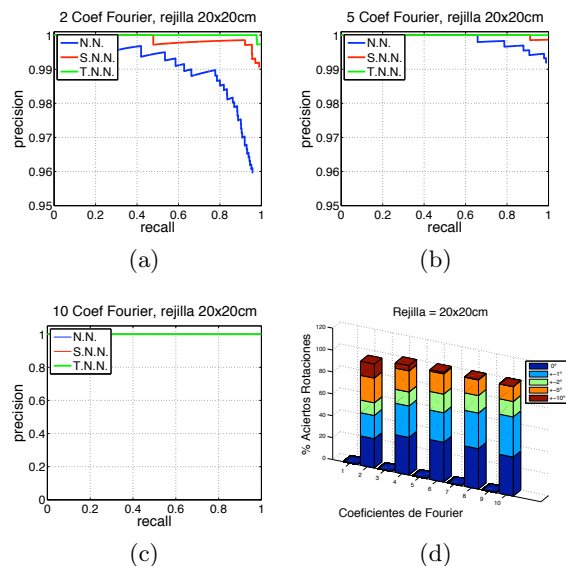


Figura 4: (a), (b), (c) Gráficas Recall-Precision con el vecino más cercano (N.N.), segundo vecino más cercano (S.N.N.) y tercer vecino más cercano (T.N.N.) y (d) precisión del desfase usando la Firma de Fourier, variando el número de componentes seleccionados por fila.

tiene bastante constante cuando se toman 8 o más componentes. Pues apreciarse que con tan sólo 2 componentes (fig 4.(a)), se tiene un 96 por ciento de precisión cuando se estudia el vecino más cercano, y casi el 100 por cien cuando se toman los tres más cercanos.

4.2. PCA SOBRE LA FIRMA DE FOURIER

Después de aplicar PCA sobre la matriz de módulos obtenida de la Firma de Fourier, se obtiene otra matriz que contiene los vectores principales seleccionados, y la proyección de las imágenes de las imágenes sobre el espacio que forman dichos vectores. Por otro lado, se mantiene la matriz de fases obtenida de la Firma de Fourier directamente, sin aplicar PCA. Para conocer dónde está el robot, primero se calcula la Firma de Fourier de la imagen de entrada. Después de seleccionar el número de coeficientes correspondiente de cada fila, se proyecta el vector de módulos en el espacio calculado con PCA. La imagen más similar en

el mapa se obtiene hallando la mínima distancia Euclídea entre la proyección de los coeficientes de la nueva imagen y los del mapa. Cuando se conoce la posición, la fase se calcula de la misma forma que cuando no se aplica PCA, pues la matriz de fases no utiliza la proyección en el nuevo espacio. La tabla 3 presenta el tiempo que necesita el algoritmo para calcular la pose tras leer la imagen. Para medir dicho tiempo, el número de vectores principales se mantiene constante. Los resultados muestran que el tiempo empleado crece con el número de imágenes que forman el mapa o cuando se seleccionan más componentes de la Firma de Fourier. Puede destacarse que, excepto en el caso de la cuadrícula de 10x10cm, el número de coeficientes que seleccionamos es más influyente que el número de imágenes que forman el mapa, ya que el cálculo de la fase sigue siendo un proceso con mucho mayor coste computacional que la localización, y esa parte del cálculo de la pose depende únicamente del número de coeficientes de Fourier que tomamos.

Tabla 3: Tiempo empleado en la estimación de la pose usando PCA sobre la Firma de Fourier seleccionando 40 vectores principales y con número de coeficientes de Fourier variable.

GRID	2Com	5Com	10Com	20Com
10x10	0.1113	0.2420	0.2844	0.3362
20x20	0.0355	0.0613	0.0985	0.1462
30x30	0.0233	0.0483	0.0838	0.1321
40x40	0.0205	0.0459	0.0815	0.1294

Tabla 4: Tiempo empleado en la estimación de la pose usando PCA sobre la Firma de Fourier seleccionando 40 coeficientes de la Transformada de Fourier de cada fila y con número de vectores principales variable.

GRID	2Vect	5Vect	10Vect	20Vect
10x10	0.1505	0.1492	0.1497	0.1497
20x20	0.1294	0.1293	0.1300	0.1295
30x30	0.1278	0.1280	0.1271	0.1273
40x40	0.1272	0.1273	0.1272	0.1270

La tabla 4 muestra el tiempo empleado cuando se aplica PCA sobre la Firma de Fourier, pero en este caso variando el número de vectores principales que se usan para proyectar las imágenes. Se puede apreciar de nuevo que el número de vectores no es muy influyente ya que el tiempo empleado por el algoritmo no cambia significativamente. Además, también se pueden ver los efectos de aplicar PCA en que no hay excesiva variación en el tiempo cuando la cuadrícula del mapa varía, pues proyectar el vector de características en el nuevo espacio es un proceso rápido y no hay diferencias significa-

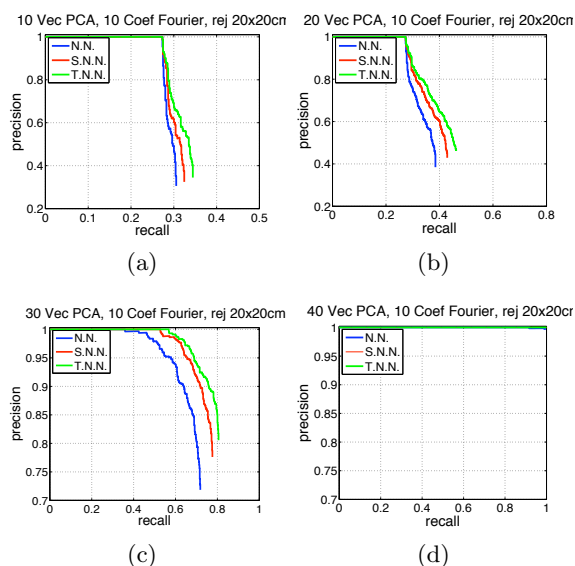


Figura 5: Gráficas Recall-Precision con el vecino más cercano (N.N.), segundo vecino más cercano (S.N.N.) y tercer vecino más cercano (T.N.N) usando PCA sobre la Firma de Fourier variando el número de vectores PCA.

tivas cuando se incrementa el número de imágenes o se usan más vectores que forman dicho espacio. En la fig.5 se aprecia que si se busca una precisión de localización relativamente alta, se requiere un elevado número de los vectores principales obtenidos tras la descomposición PCA, lo que significa perder las ventajas de aplicar PCA. Además, en la mayoría de los experimentos, el número de coeficientes de Fourier necesarios es mayor que cuando únicamente se emplea la Firma de Fourier, incrementando la memoria requerida. Los resultados del cálculo del desfase no se incluyen porque, como se ha comentado anteriormente, son exactamente los mismos que los mostrados en fig. 1(d) ya que su cálculo no varía.

4.3. HISTOGRAMA DE ORIENTACIÓN DEL GRADIENTE

Cuando llega una nueva imagen, se calcula el histograma de orientación del gradiente usando celdas con el mismo tamaño que las usadas para construir el mapa. El tiempo requerido para encontrar la pose del robot varía dependiendo tanto de las celdas horizontales como verticales. Para localizar al robot, se usa la información de las celdas horizontales, mientras que para el cálculo de la fase necesitamos la de las celdas verticales. En ambos casos, los resultados se obtienen calculando la mínima distancia Euclídea entre los histogramas de la nueva imagen y los de las imágenes del mapa.

En la localización, como no se superponen las cel-

Tabla 5: Tiempo empleado en la estimación de la pose usando HOG variando la altura de las celdas horizontales.

GRID	2 Pix	5 Pix	10 Pix	20 Pix
10x10	0.1113	0.2420	0.2844	0.3362
20x20	0.0355	0.0613	0.0985	0.1462
30x30	0.0233	0.0483	0.0838	0.1321
40x40	0.0205	0.0459	0.0815	0.1294

Tabla 6: Tiempo empleado en la estimación de la pose usando HOG variando la distancia de aplicación de las celdas verticales.

GRID	2 Pix	5 Pix	10 Pix	20 Pix
10x10	0.6599	0.5259	0.4979	0.4857
20x20	0.2777	0.1598	0.1240	0.1136
30x30	0.2546	0.1301	0.0982	0.0866
40x40	0.2497	0.1247	0.0943	0.0822

das, cuando el tamaño de estas es menor necesitamos usar un mayor número, lo que supone calcular más histogramas. Por ello, cuando se reduce la altura de las celdas horizontales, el tiempo empleado es mayor (tabla 5). Además, cuando el mapa se forma con más imágenes, se tienen que realizar más comparaciones para encontrar la imagen más cercana, lo que significa emplear más tiempo. Pero con pocas celdas, aumenta la dificultad de reconocer la posición correcta, tal y como se muestra en la fig. 6. En la tabla 6 se comprueba el efecto de variar la distancia entre las celdas verticales. Este parámetro condiciona el cálculo del desfase. Cuanto mayor sea la distancia de aplicación de las celdas, menos tiempo se requiere, pues el número de histogramas es menor. Sin embargo, para mejorar la precisión del ángulo debemos reducir dicha distancia.

5. CONCLUSIONES

Este trabajo se ha centrado en la comparación de diferentes algoritmos basados en apariencia aplicados a la creación de un mapa denso de un entorno real, usando imágenes omnidireccionales. Se han presentado tres métodos diferentes para comprimir la información del mapa. Todos ellos han demostrado ser válidos para llevar a cabo la estimación de la pose de un robot dentro de un mapa. A la vista del estudio, la firma de Fourier demuestra ser el método más eficiente ya que, seleccionando pocos componentes de la transformada de Fourier de cada fila, se obtienen buenos resultados. Por ello, el mapa no necesita gran cantidad de memoria para ser almacenado, y la comparación no es un proceso con alto coste computacional. Sin embargo, la aplicación de PCA a la firma de Fourier no ha tenido las ventajas que se

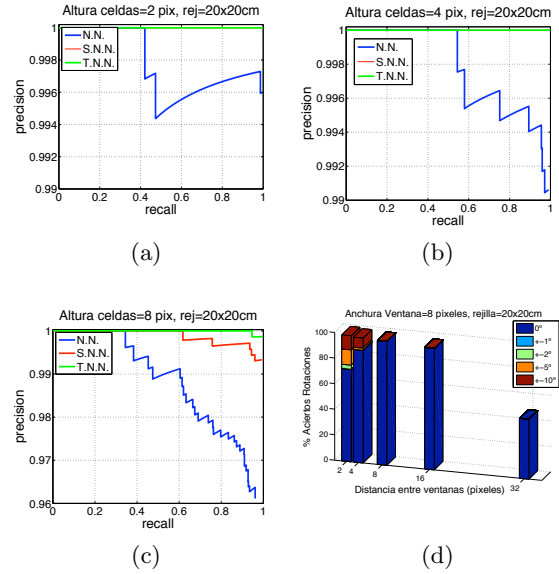


Figura 6: (a), (b), (c) Gráficas Recall-Precision con el vecino más cercano (N.N.), segundo vecino más cercano (S.N.N.) y tercer vecino más cercano (T.N.N.) y (d) precisión de la fase usando HOG variando la altura de las celdas horizontales.

esperaba, pues crear el espacio de proyección es un proceso computacional pesado, tardando una media de tres veces más que el primer algoritmo. Además, para tener buenos resultados se necesita mantener la gran mayoría de vectores principales obtenidos, y seleccionar más coeficientes de Fourier, por lo que el tamaño del mapa no se reduce, si no que, al contrario, la memoria requerida es el triple en el mejor de los casos. En ambos casos, la precisión de la orientación depende únicamente del número de componentes de Fourier, y el error en su estimación es igual o menor a 5 grados en la gran mayoría de los experimentos. Con respecto a HOG, los resultados demuestran que es un método robusto, aunque no es muy flexible debido a sus requerimientos de tiempo y memoria, pues el cálculo de los histogramas es una tarea que consume relativamente mucho tiempo. Para llevar a cabo la localización, el tiempo es aproximadamente 10 veces mayor que en el caso del algoritmo de Fourier, aunque varía significativamente si modificamos la cuadrícula de imágenes que forman el mapa. En cuanto la memoria, se sitúa entre la Firma de Fourier y PCA sobre Fourier, siendo aproximadamente un cincuenta por ciento mayor que el primer método. La parte menos efectiva de este algoritmo está en el cálculo de la fase, ya que los grados están muestreados en función del número de ventanas verticales que usamos. Si las limitaciones de memoria son más importantes que las de tiempo, y no se requiere alta precisión en la orientación, puede ser un método interesante.

Este artículo muestra de nuevo el amplio rango de posibilidades de los métodos basados en apariencia global aplicados a robots móviles, y sus resultados animan a continuar con su estudio en profundidad.

Agradecimientos

Este trabajo ha sido apoyado por el gobierno español a través del proyecto DPI2007-61197 “Sistemas de percepción visual móvil y cooperativo como soporte para la realización de tareas con redes de robots”.

Referencias

- [1] M. Cummins and P. Newman. Fab-map: Probabilistic localization and mapping in the space of appearance. In *International Journal of Robotics Research* 27 (6), pp. 647-665, 2008.
- [2] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In *Proc of the IEEE Conf on Computer Vision and Pattern Recognition, San Diego, USA. Vol. II*, pp. 886-893, 2005.
- [3] A. Gil, O. Martínez, M. Ballesta, and O. Reinoso. A comparative evaluation of interest point detectors and local descriptors for visual slam. SPRINGER Machine Vision and Applications, 2009.
- [4] M. Jogan and A. Leonardis. Robust localization using eigenspace of spinning-images. In *Proc. IEEE Workshop on Omnidirectional Vision, Hilton Head Island, USA*, pp. 37-44. IEEE, 2000.
- [5] M. Kirby. *Geometric data analysis*. Wiley Interscience, 2000.
- [6] B. Krose, R. Bunschoten, S. Hagen, B. Terwijn, and N. Vlassis. Visual homing in environments with anisotropic landmark distribution. In *Autonomous Robots*, 23(3), 2007, pp. 231-245, 2007.
- [7] E. Menegatti, T. Maeda, and H. Ishiguro. Image-based memory for robot navigation using properties of omnidirectional images. In *Robotics and Autonomous Systems. Vol. 47, No. 4*, pp. 251-276, 2004.
- [8] R. Moeller, A. Vardy, S. Kreft, and S. Ruwisch. Visual homing in environments with anisotropic landmark distribution. In *Autonomous Robots*, 23(3), 2007, pp. 231-245, 2007.
- [9] L. Paya, L. Fernández, O. Reinoso, A. Gil, and D. Ubeda. Appearance-based dense maps creation. In *6th Int Conf on Informatics in Control, Automation and Robotics ICINCO 2009*. Ed. INSTICC PRESS - pp. 250-255, 2009.
- [10] P. E. Rybski, S. Roumeliotis, M. Gini, and N. Papanikopoulos. Appearance-based mapping using minimalistic sensor models. In *Autonomous Robots, vol. 24, no. 3*, pp. 229-246. Springer Netherlands, 2008.
- [11] D. Scaramuzza, R. Siegwart, and A. Martinelli. A robust descriptor for tracking vertical lines in omnidirectional images and its use in mobile robotics. In *International Journal of Robotics Research* 28 (2), pp. 149-171, 2009.
- [12] C. Siagian and L. Itti. Rapid biologically-inspired scene classification using features shared with visual attention. In *IEEE Trans. Pattern Analysis and Machine Intelligence* 29 (2), pp. 300-312, Feb 2007, 2007.
- [13] A. Stimec, M. Jogan, and A. Leonardis. Unsupervised learning of a hierarchy of topological maps using omnidirectional images. In *International Journal of Pattern Recognition and Artificial Intelligence*. 22(4), pp 639-665, 2008.
- [14] S. Thrun. Robotic mapping: A survey, in exploring artificial intelligence. In *The New Millennium*, pp. 1-35. Morgan Kaufmann Publishers, San Francisco, USA, 2003.
- [15] Z. Zivkovic, O. Booij, and B. Kruse. From images to rooms. In *Robotics and Autonomous Systems*. 55(5), pp. 411-418, 2007.